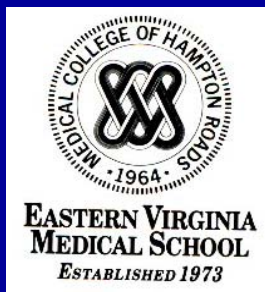


# Signal and Noise: Garbage in Garbage out

O. John Semmes

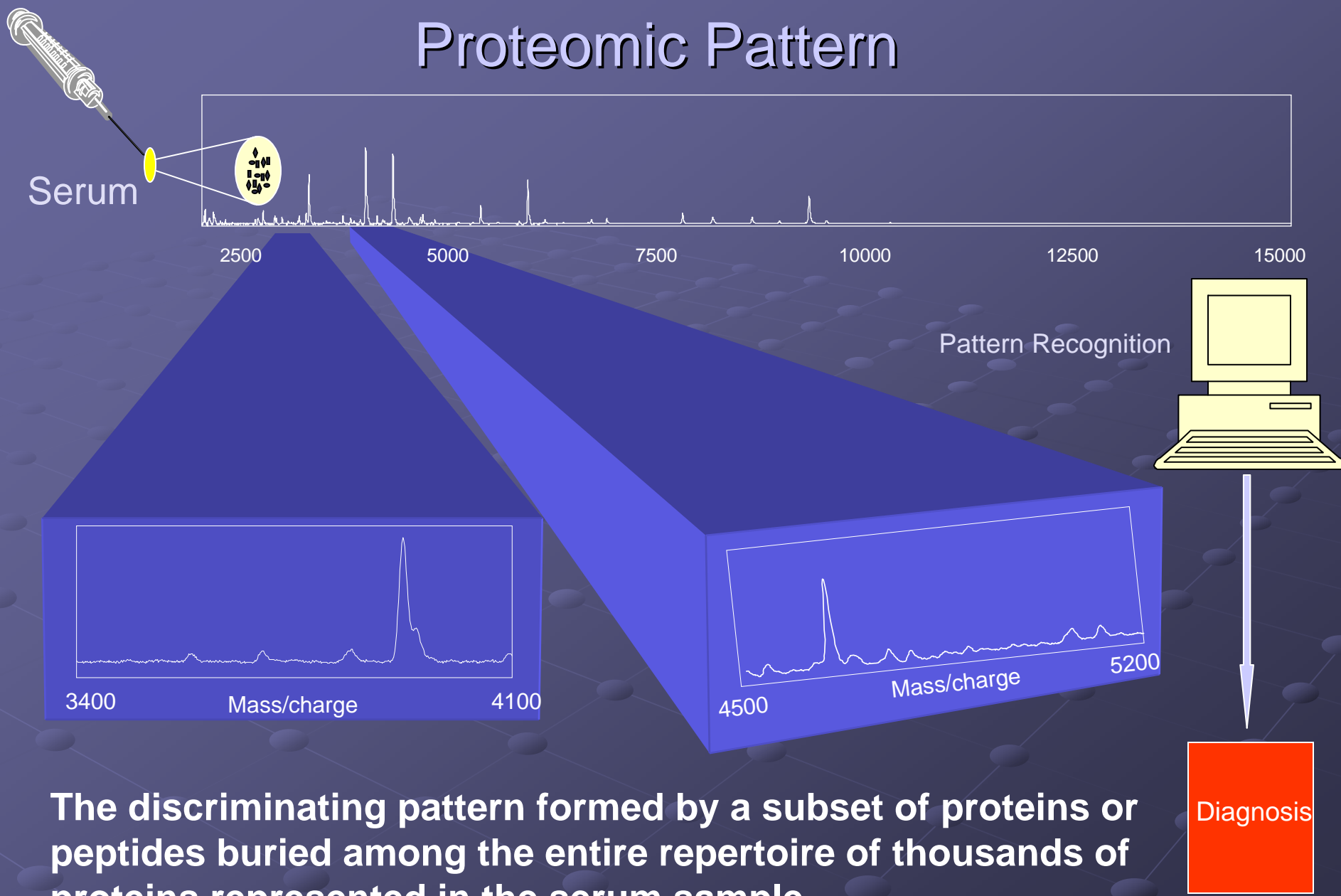
*Director, Center for Biomedical Proteomics  
Department of Microbiology and Molecular Cell Biology  
Virginia Prostate Center  
Eastern Virginia Medical School Norfolk, VA*



Eastern Virginia Medical School  
Discovery Laboratory



# Proteomic Pattern



**The discriminating pattern formed by a subset of proteins or peptides buried among the entire repertoire of thousands of proteins represented in the serum sample.**

# THINGS WE CAN DO TO MAKE THE PATTERNS BETTER

1. **Some things to do before the Mass Spec (TOF-tube)**
2. **Some things to do after the Mass Spec (TOF-tube)**

# Achieving Reproducibility in SELDI

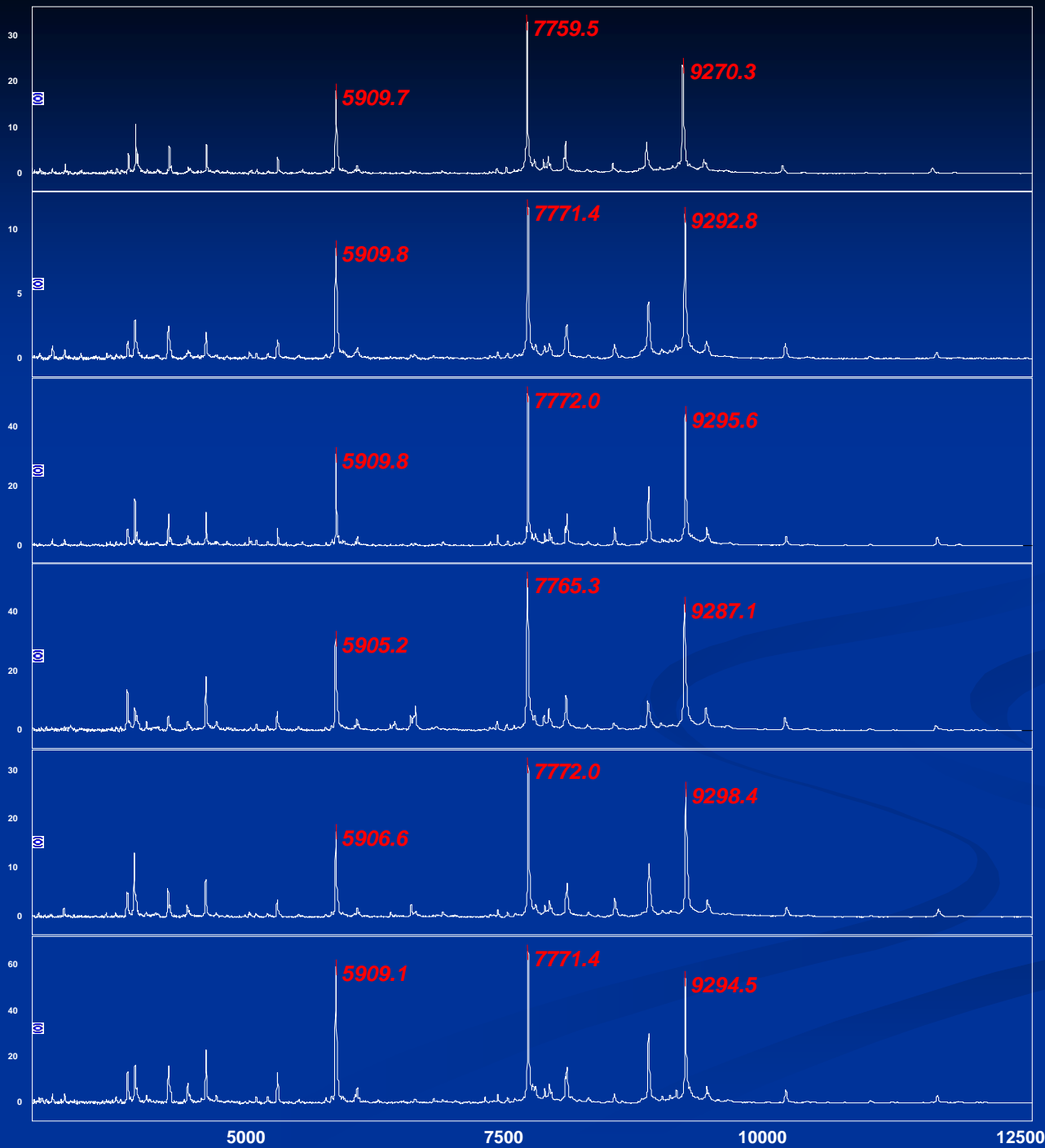
1. Get a gatekeeper for experimental design and interpretation.
2. Use “standard” samples; external (spiked) proteins, internal proteins, serum sample for QC.
3. Synchronization/optimization of instrument output using the QC sera. Laser/detector settings.
4. Constant monitoring and adjustment of parameters.
5. Automation of sample processing steps.
6. Find out what the peaks are! (robustness)

**Table 1. QC spectra criteria**

Protein	Signal to Noise Ratio (S/N)	Resolution
Insulin	N/A	600
IgG	700	N/A
Peak1: 5906 ± 0.2%	>40	>400
Peak 2: 7768 ± 0.2%	>80	>400
Peak 3: 9289 ± 0.2%	>80	>400

Quality control assessment of the PBSII is based on signal to noise ratio (s/n) and resolution. The table provides the s/n and resolution required for each QC protein used in Phase IA for a site to proceed to Phase II.

# QC standardization At six Validation Sites



EVMS

UAB

UPCI

CPDR

JHMI

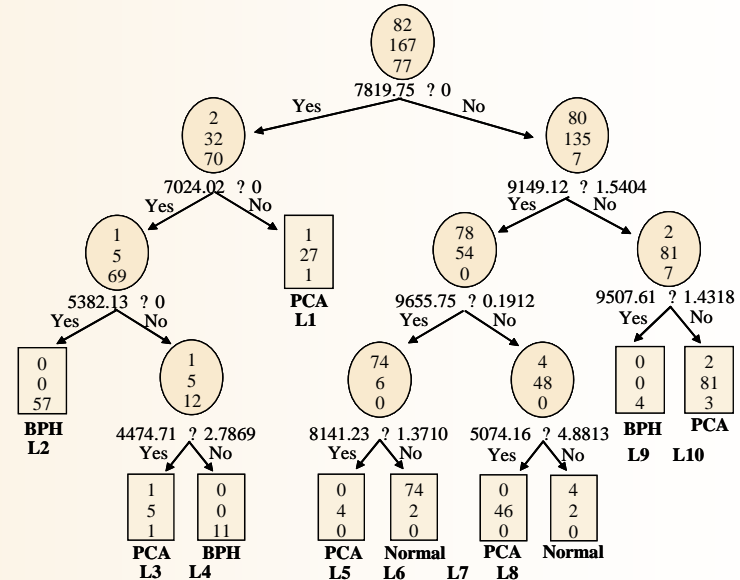
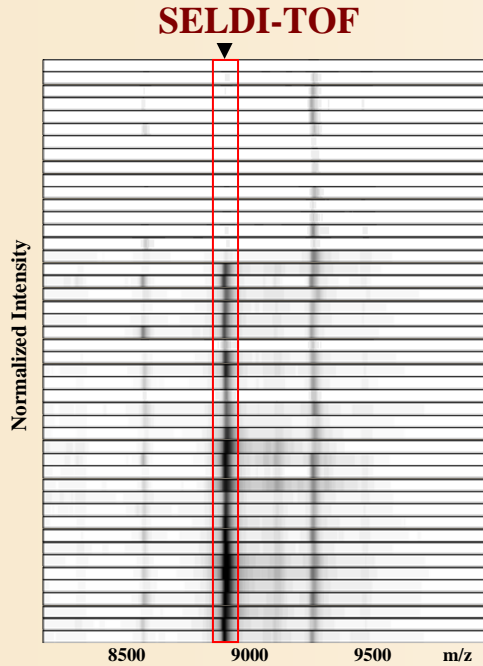
UTHSCSA

**Table 2b Inter-Lab variability**

		<b>Mass</b>	<b>Intensity</b>	<b>S/N</b>	<b>Resolution</b>
<i>Peak 1</i>	<b>average</b>	5906.47	26.57	163.06	460.73
	<b>stdev</b>	6.70	9.67		107.72
	<b>CV</b>	<b>0.0011</b>	<b>0.36</b>		<b>0.23</b>
<i>Peak 2</i>	<b>average</b>	7768.61	35.94	242.75	505.54
	<b>stdev</b>	8.41	6.25		82.77
	<b>CV</b>	<b>0.0010</b>	<b>0.17</b>		<b>0.16</b>
<i>Peak 3</i>	<b>average</b>	9289.18	30.96	244.03	439.28
	<b>stdev</b>	9.89	4.70		77.35
	<b>CV</b>	<b>0.0011</b>	<b>0.15</b>		<b>0.18</b>

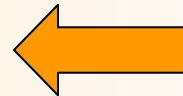
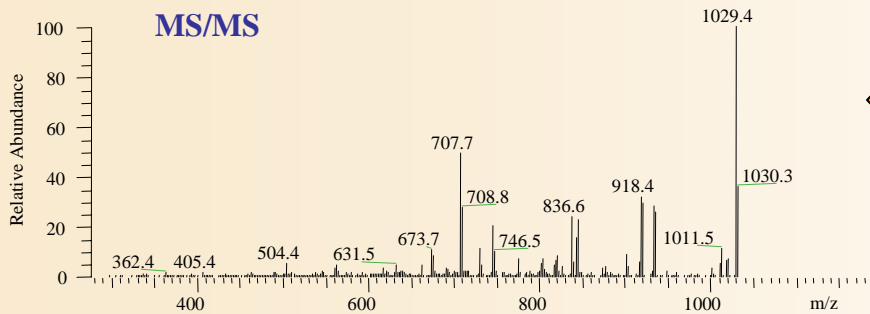
# Summary of Biomarker Discovery and Identification

## Classification and Regression Tree Analysis

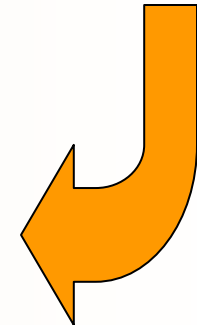
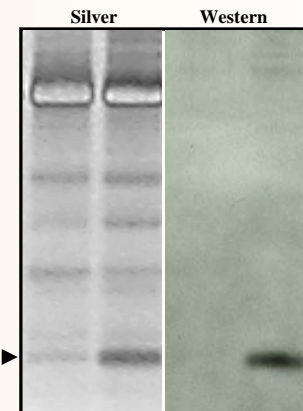


[Cancer Research 62, 3609-3614, July 1, 2002] © 2002  
American Association for Cancer Research

## Identification



## Purification





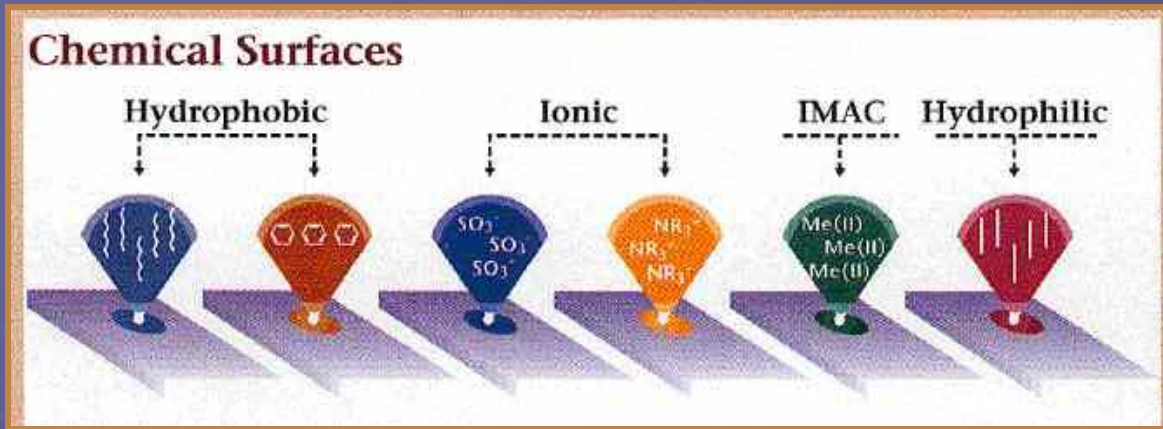
Things to do after the tube

## ***Analysis of Source of Variation***

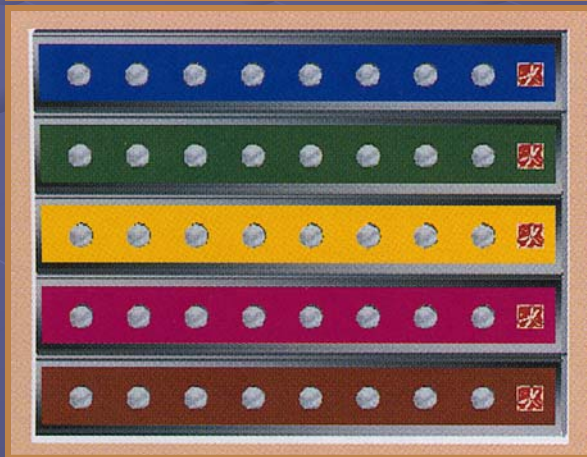
1. Metrologic Analysis of SELDI-TOF Process.
2. Spectral Analysis of Output.

# Proteomics Using SELDI Technology

## Surface Enhanced Laser Desorption



← **Surface Chemistries** Each chip binds a specific set of proteins based on the chromatographic surface of the ProteinChip®.



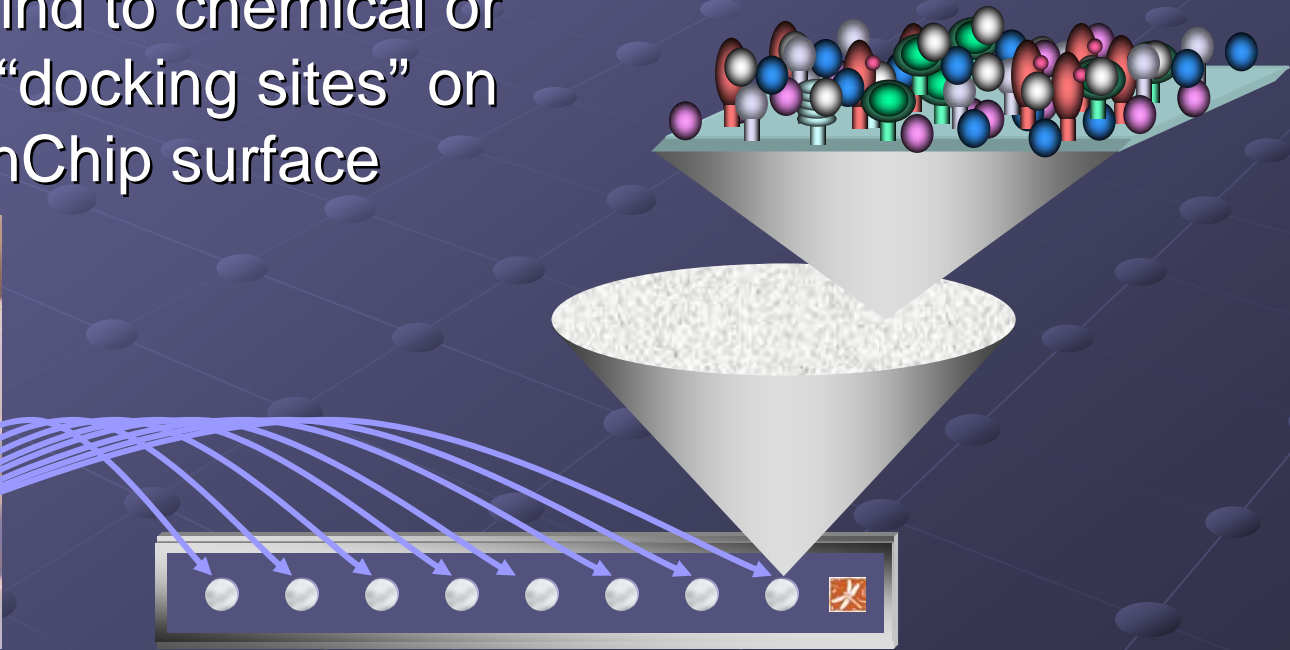
### ← **Protein Chips**

Each spot on the chip will contain sample from a control or diseased/treated source. The spots are analyzed separately and a mass spectra is created for each spot representing the proteins bound to the chip surface.



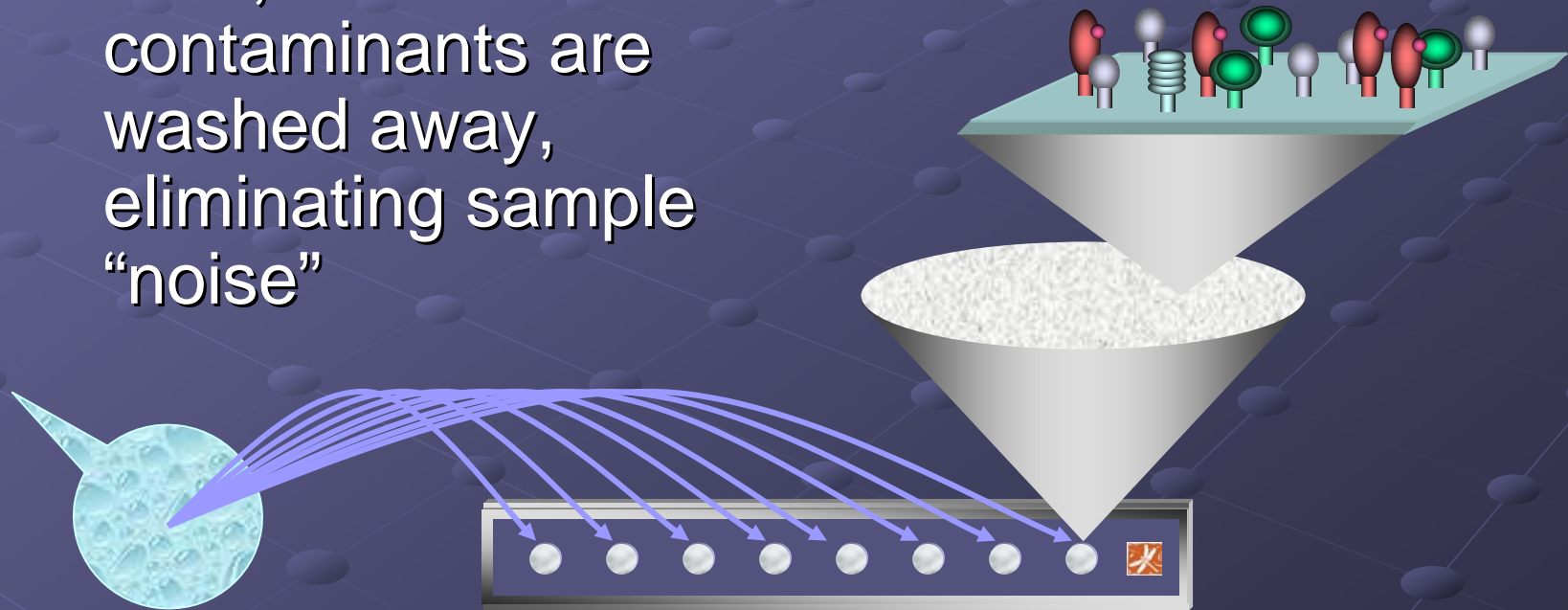
# ProteinChip Technology: Protein Binding

- Crude sample is placed (and processed) on a ProteinChip Array
- Proteins bind to chemical or biological “docking sites” on the ProteinChip surface



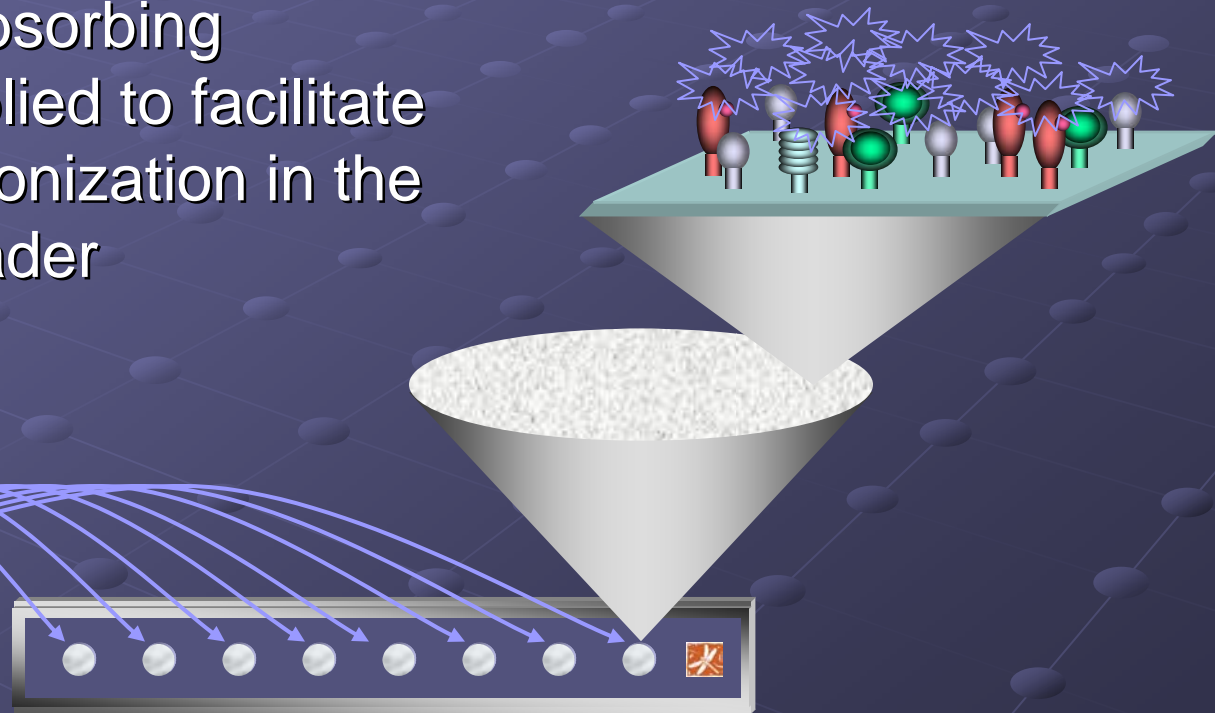
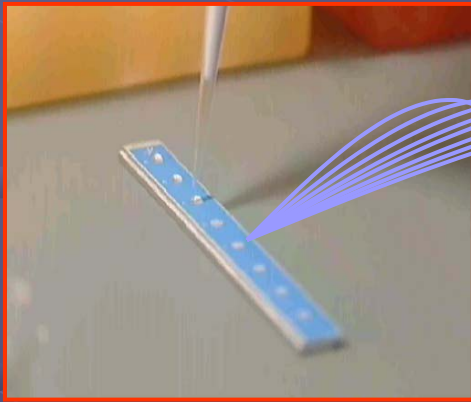
# ProteinChip Technology: Washing Reduces Non-Specific Binding

- Non-binding proteins, salts, and other contaminants are washed away, eliminating sample “noise”

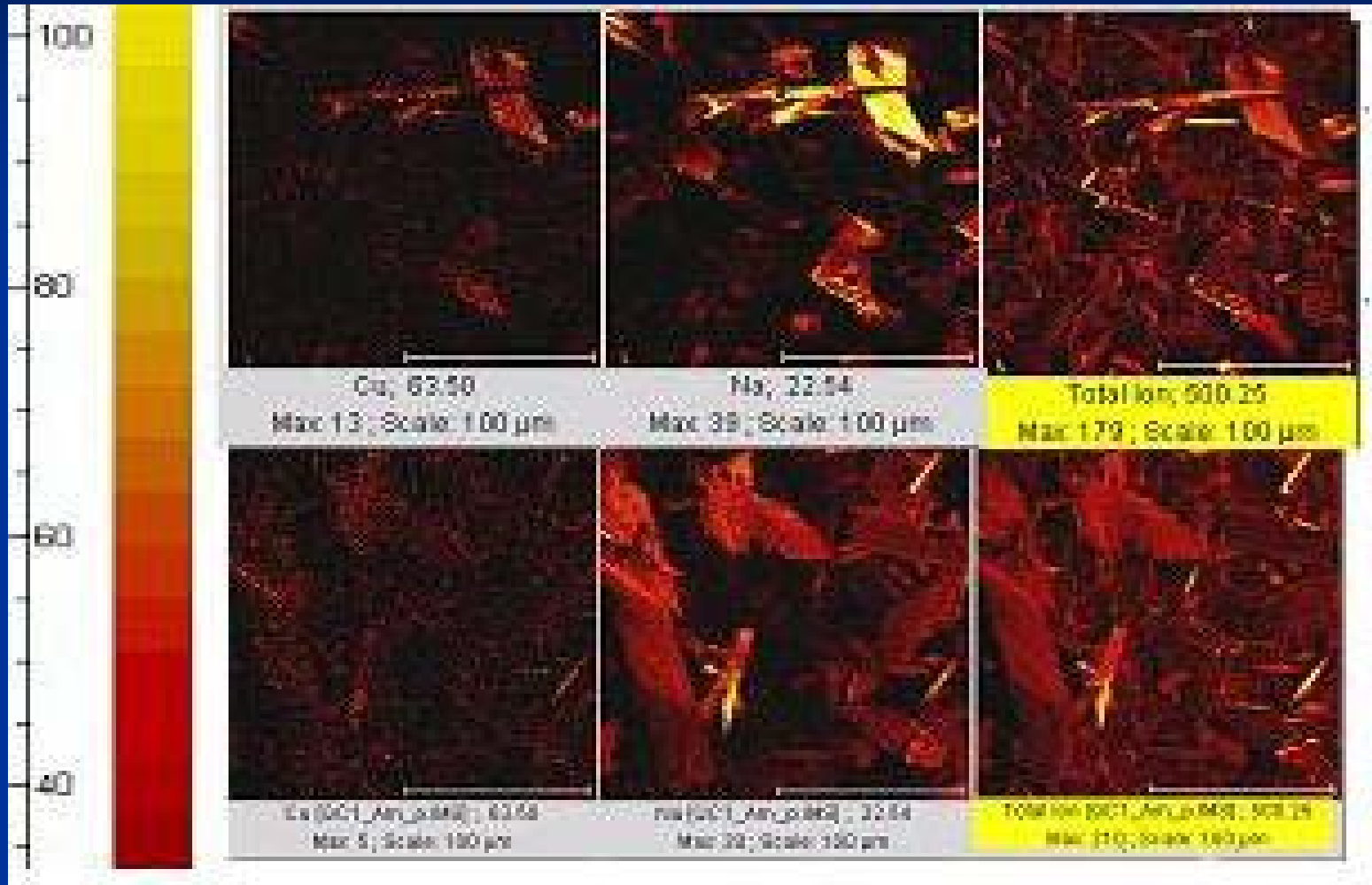


# ProteinChip Technology: Addition of EAM

- EAM (Energy Absorbing Molecule) is applied to facilitate desorption and ionization in the ProteinChip Reader



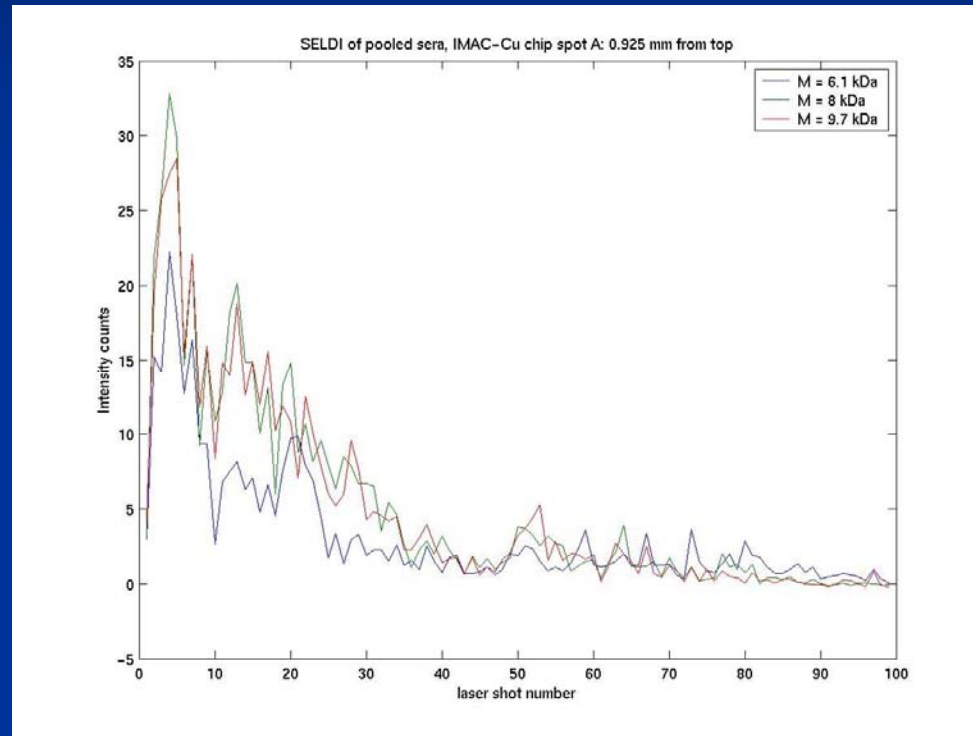
# Desorption Surface



-

+

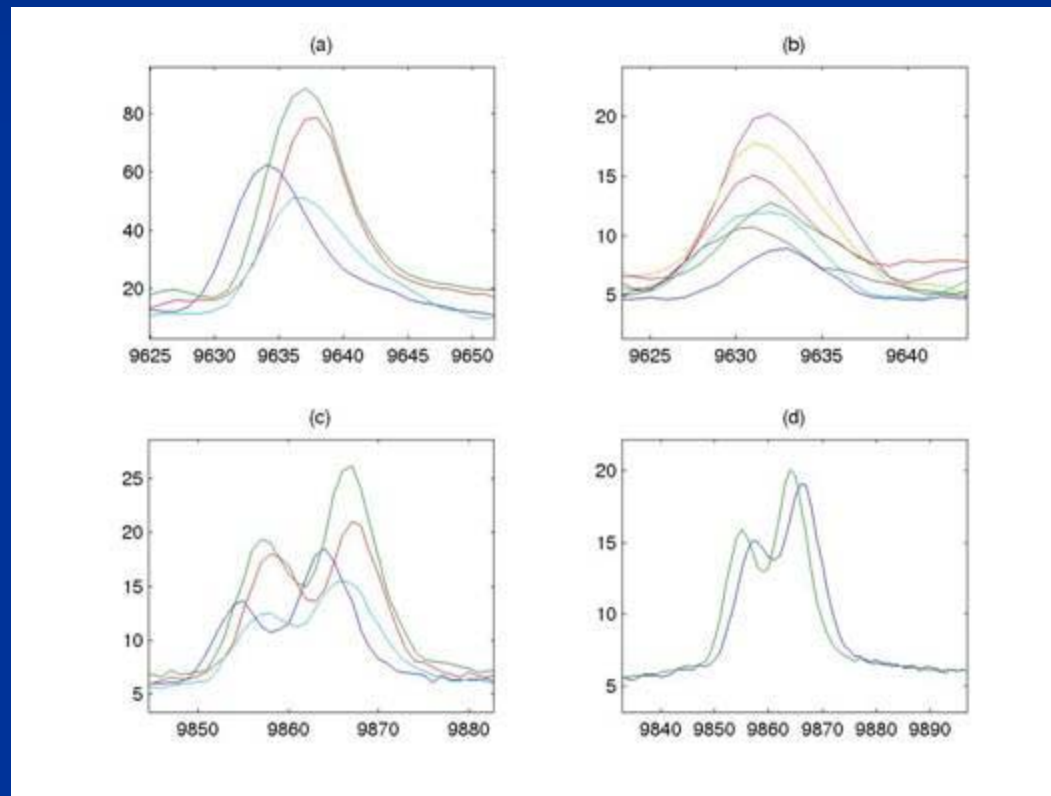
# Limited Inefficient Desorption



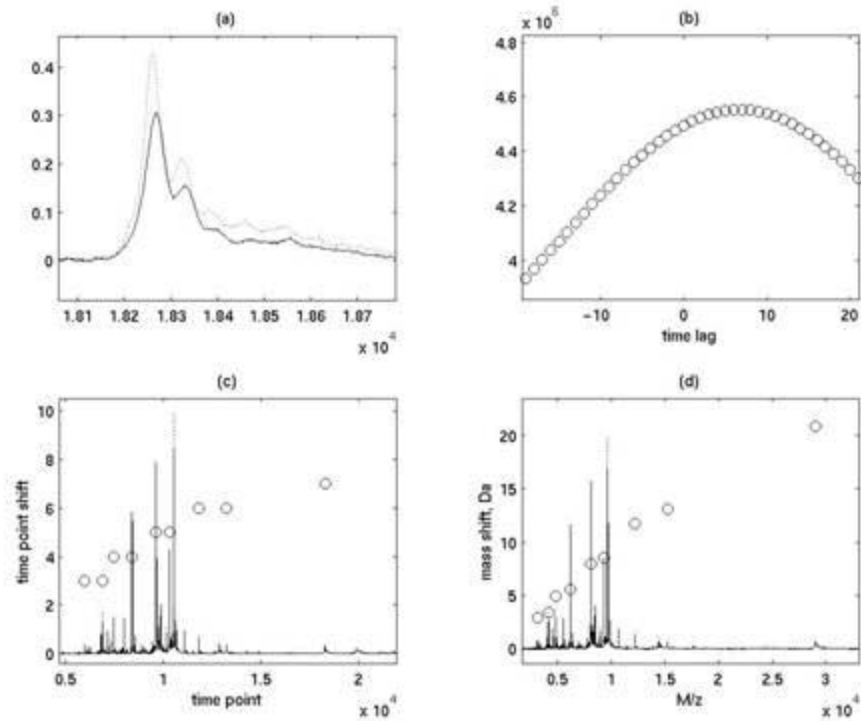
# Nano-Scale Surface Polishing



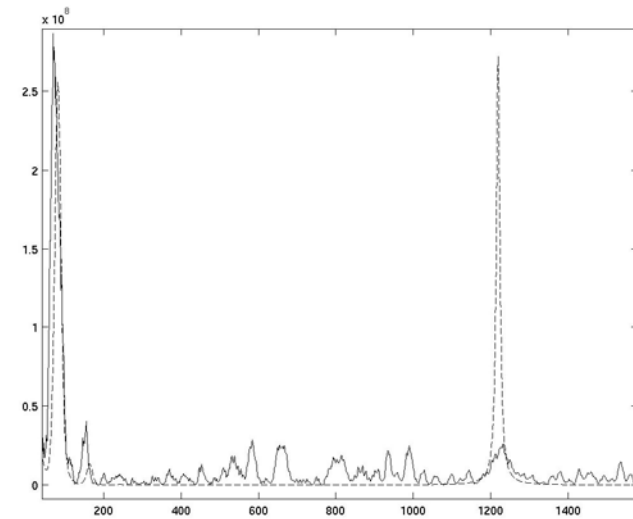
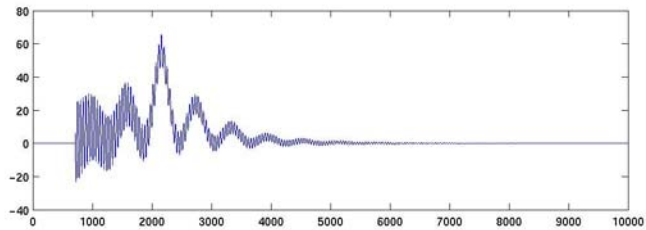
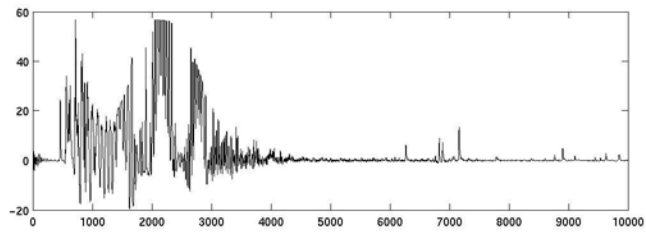
# Peak Jitter Between Single Laser Shots Reduced Resolution



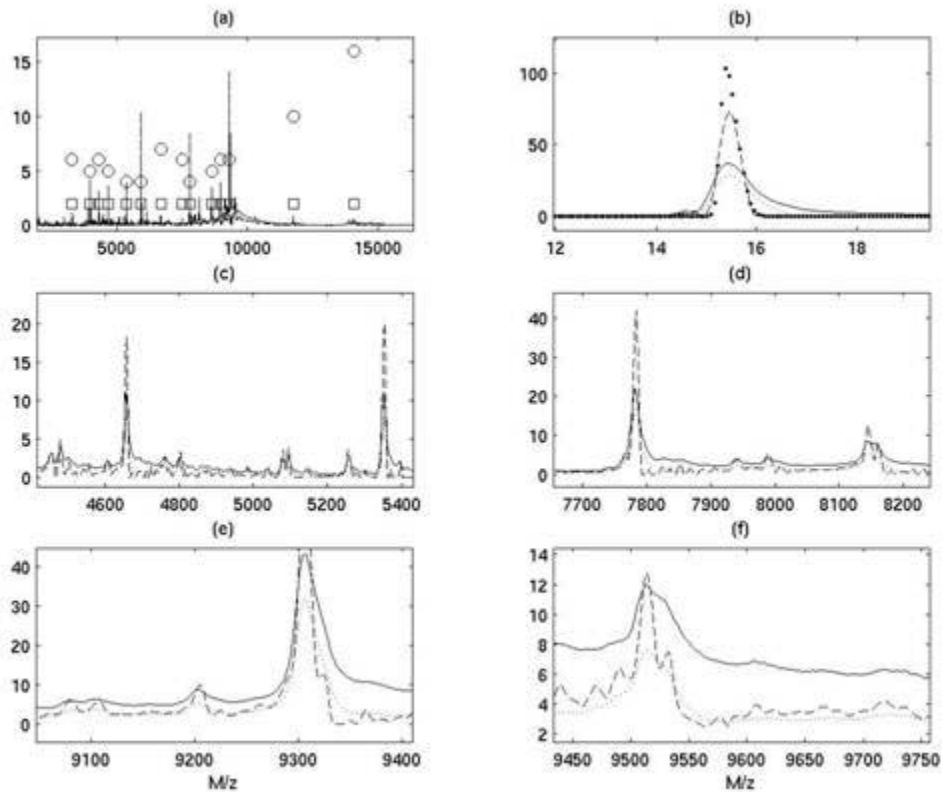
# Automatic Dejitter



# Denoising Filters

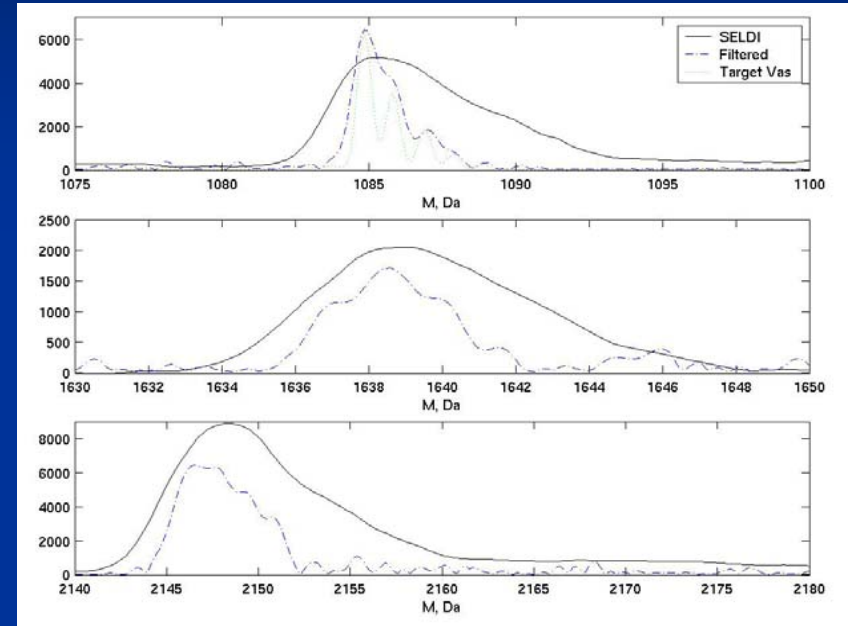
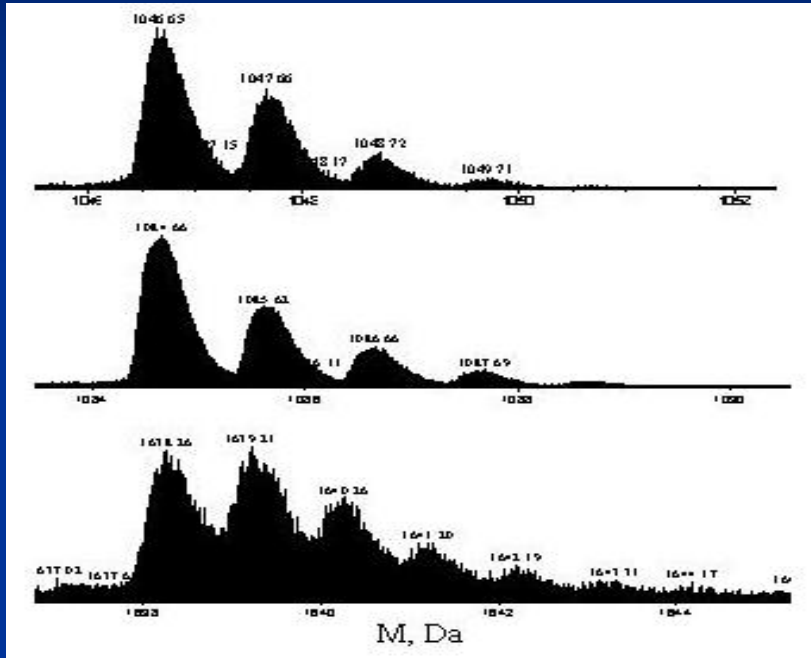


# Trace Add-Back Filters



Model based target filtering

# Trace Add-Back Filters



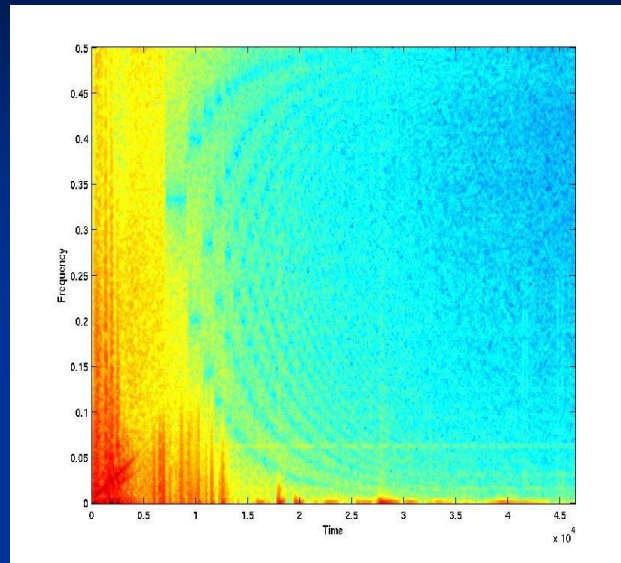
Peptide standards with SIMS-resolved isotopic structure

Best approach may involve  
Internal Standards with known isotope structure

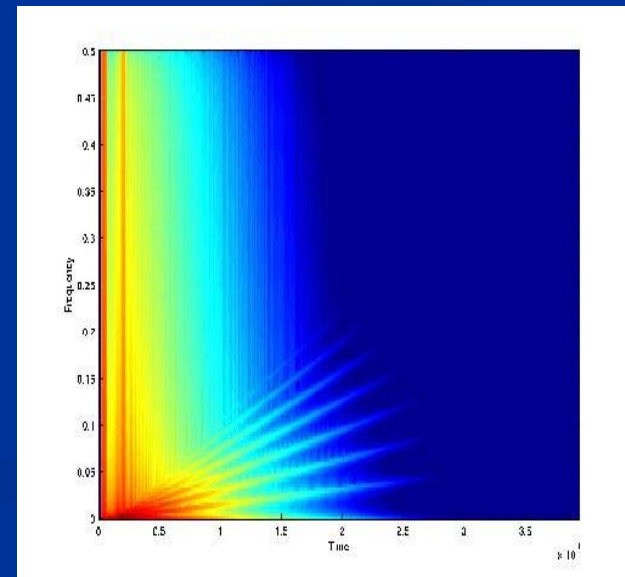
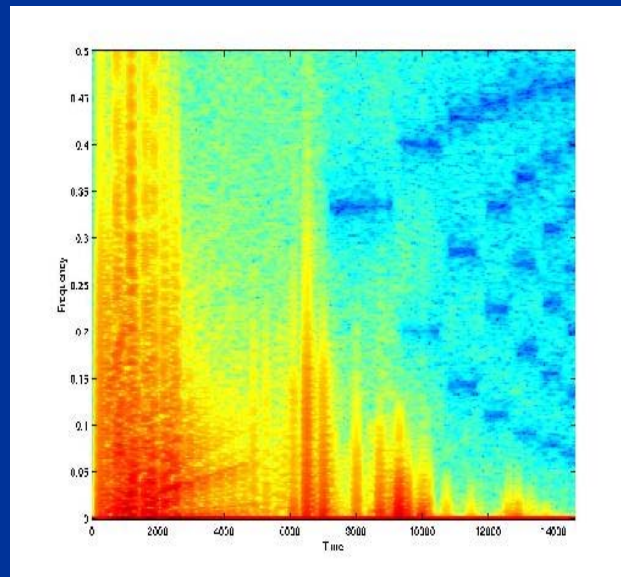
Placing external proteins in data valleys

The success of denoising filters depends on defining baseline

# Spectral Analysis

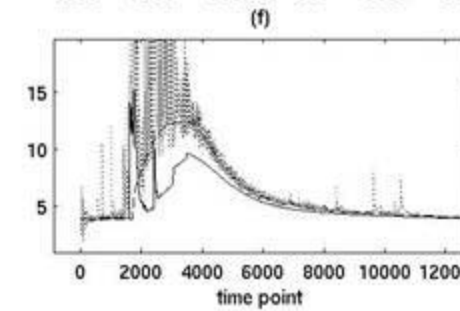
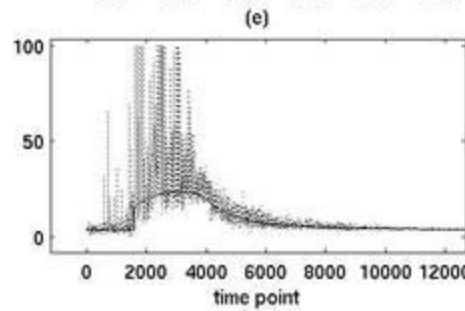
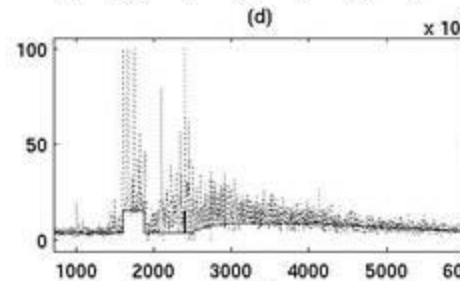
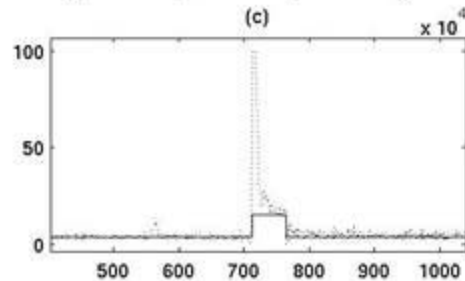
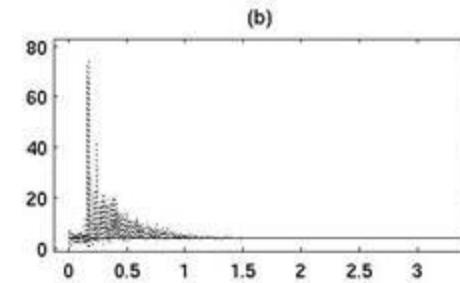
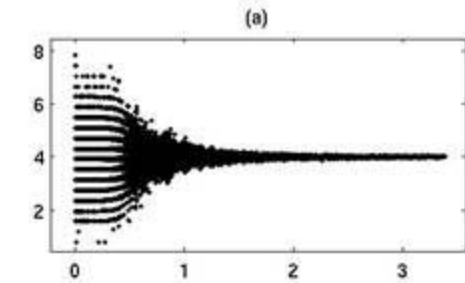


Detector Overload



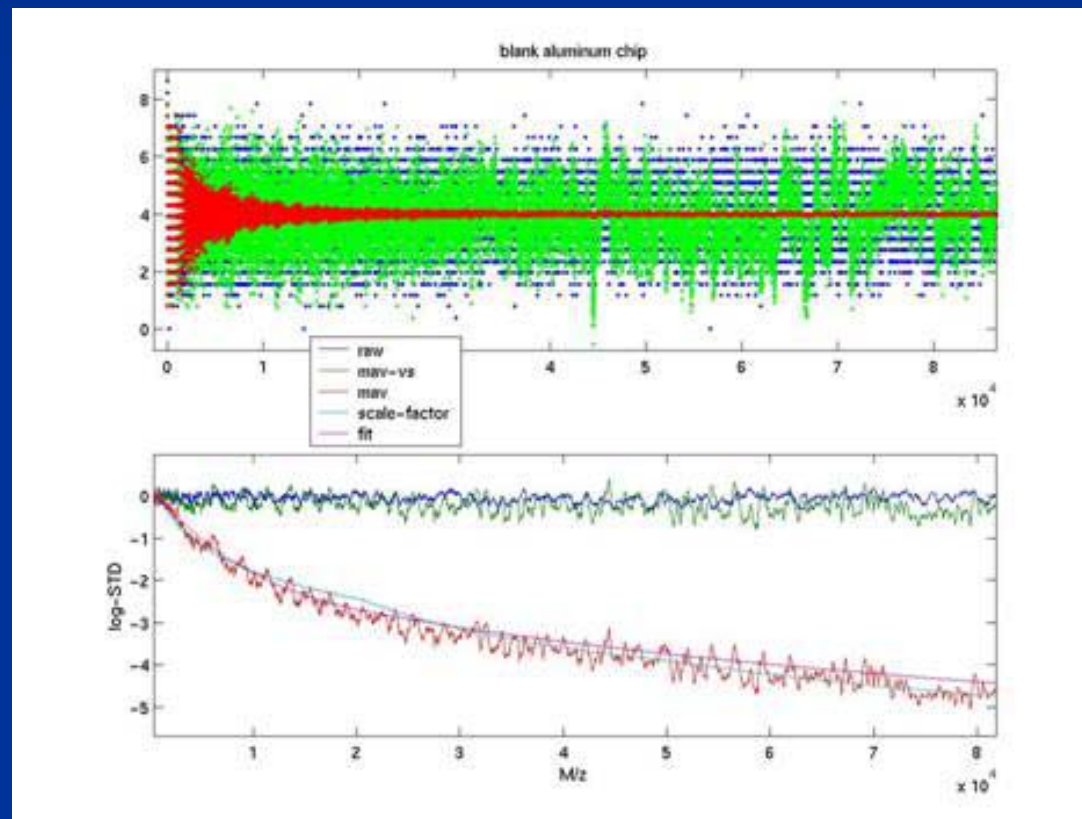


# Effect of Detector Overload On Baseline

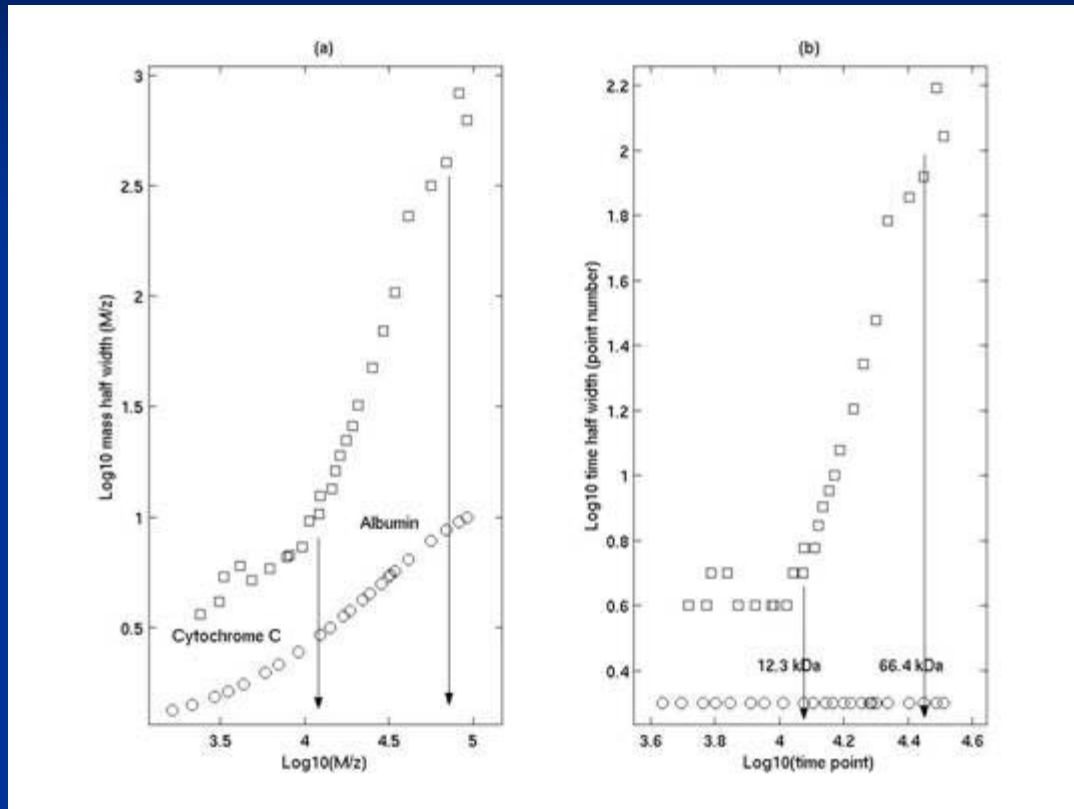


# Variance Rescaling

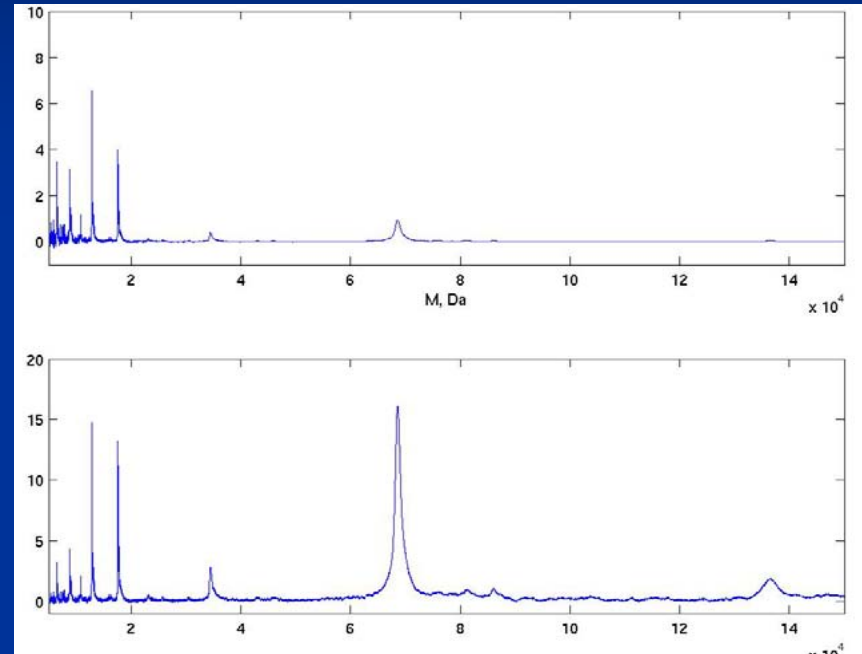
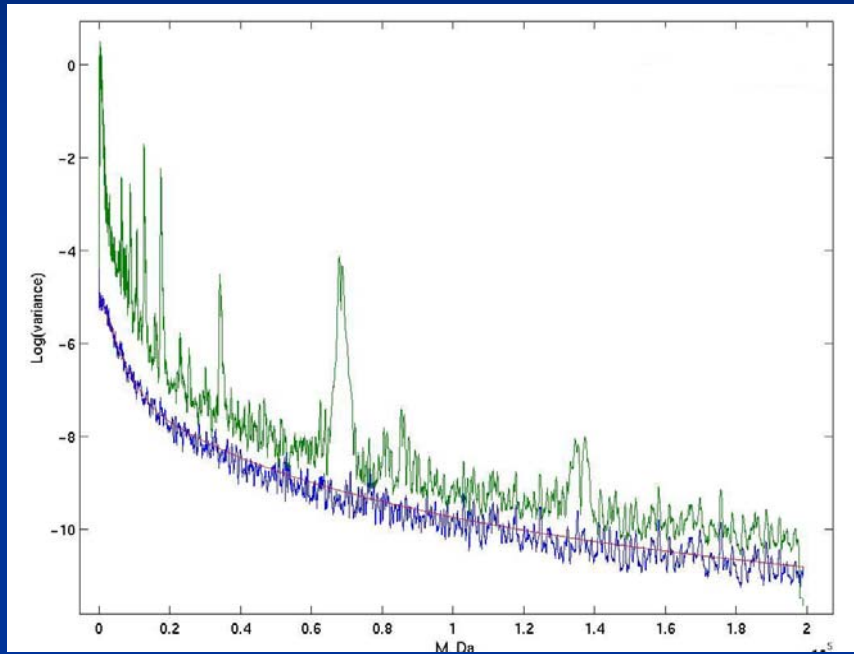
## Effect of Default Moving Average Filter



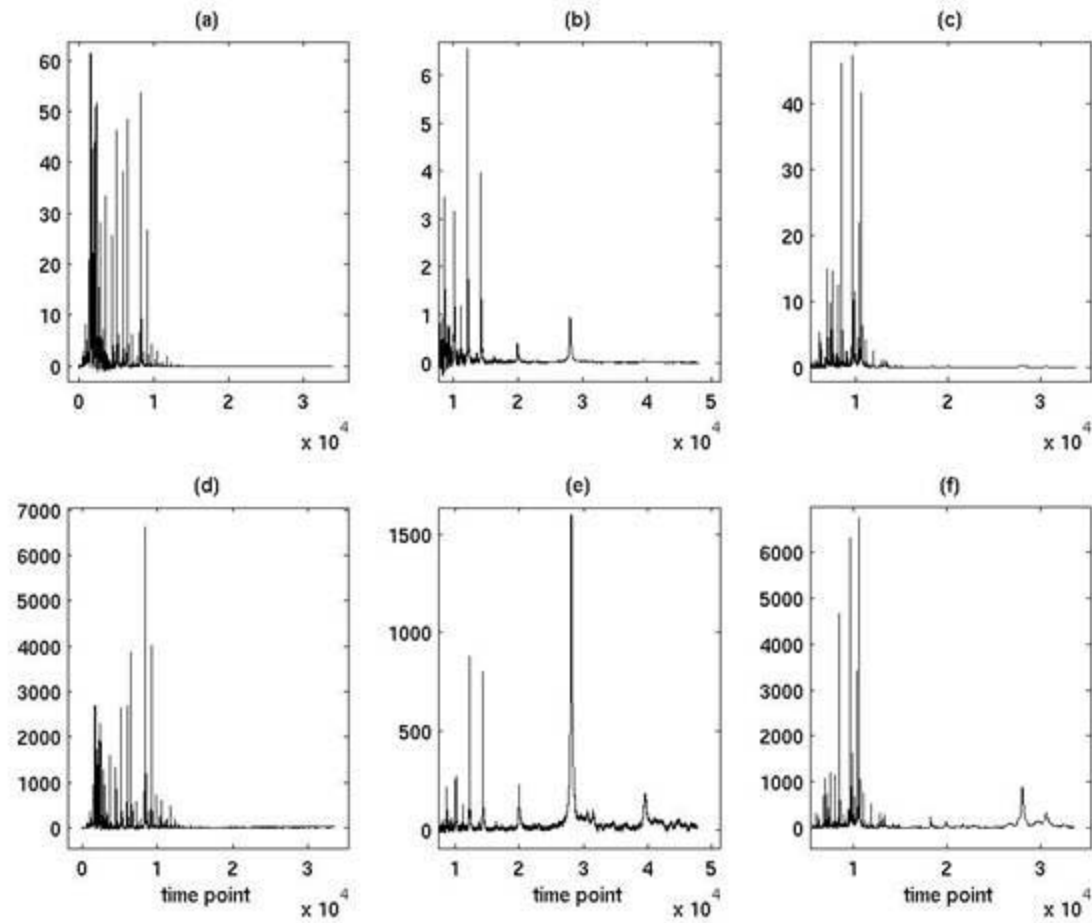
# Mass dependence of peak width and default MAV



# Mass Dependence of Variance



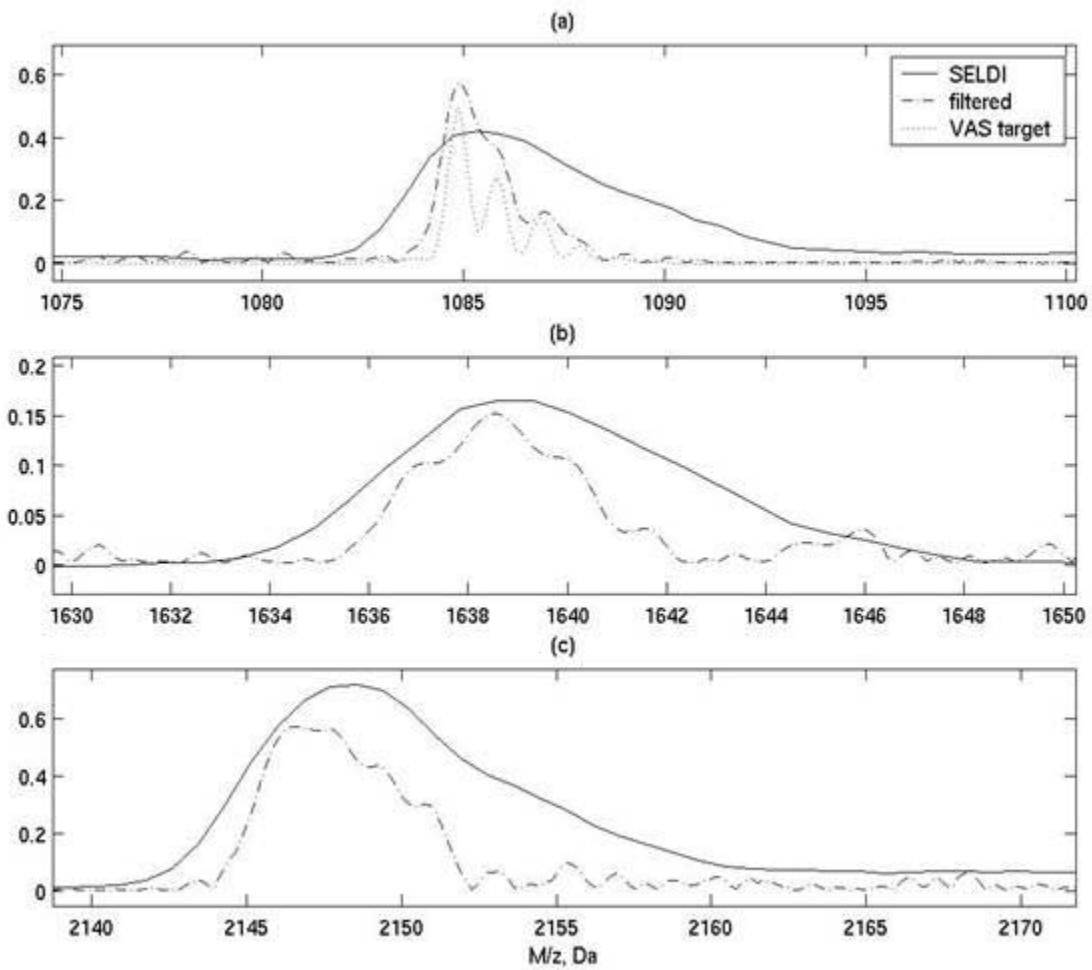
# Variance Rescaling: Stationary Noise, Increased Sensitivity



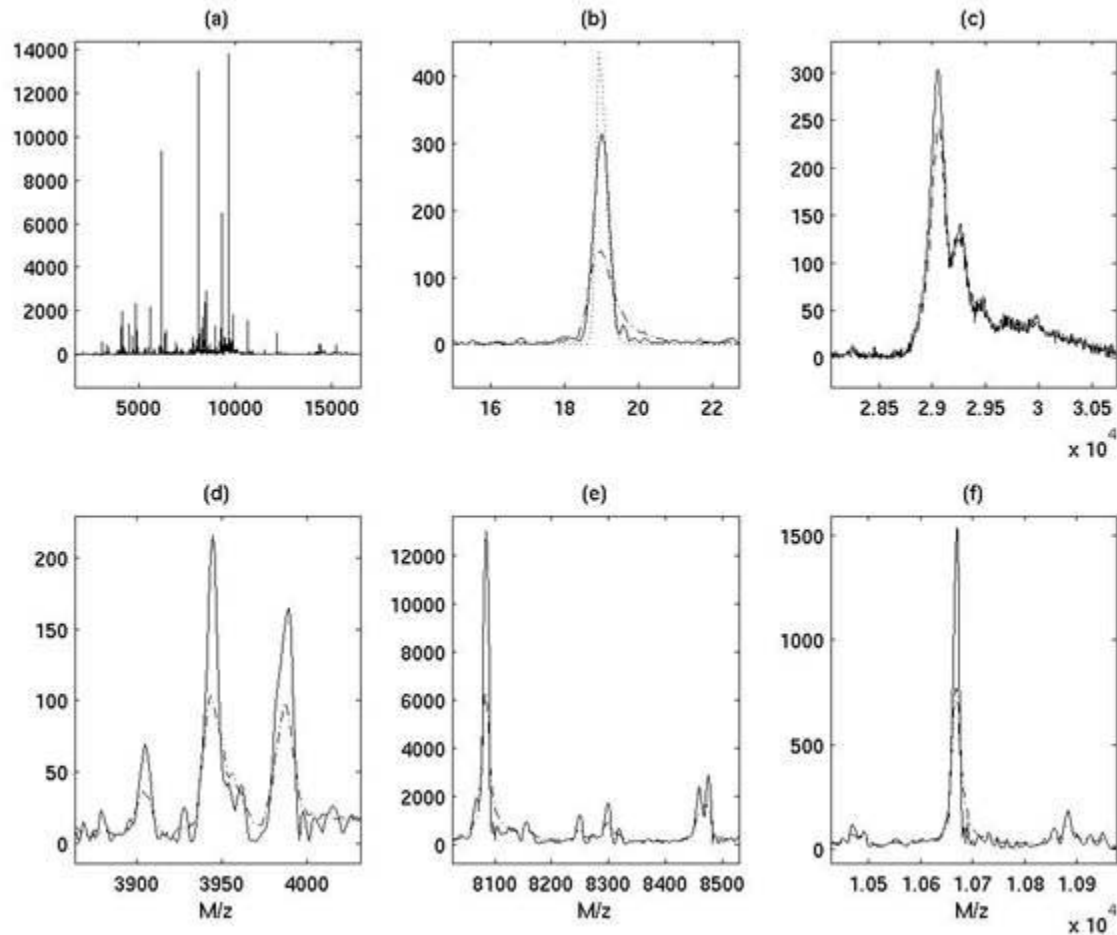
# Putting it all together



# Enhanced Resolution of Calibrant Peaks



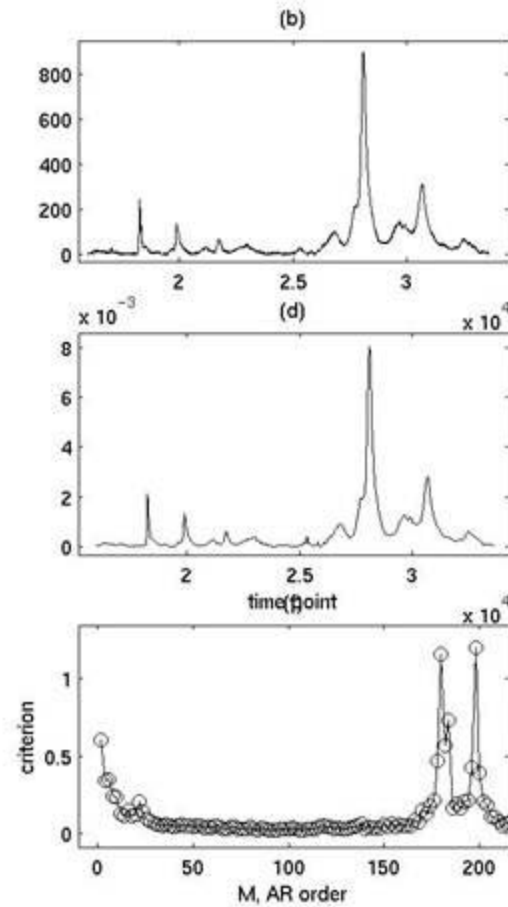
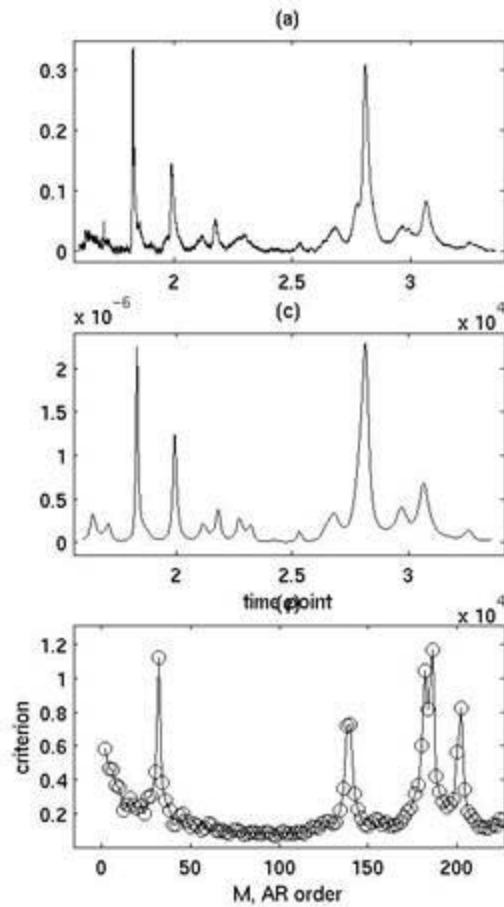
# Enhanced Resolution in Pooled Serum





# Default

# BKG-Sub, MAV, Var-Rescale



# Summary

Improving the processing of data output can dramatically improve sensitivity, resolution and reproducibility.

The Fold improvement may equal that of the “High resolution” SELDI-QStar.

Lookout for default Settings

# Eastern Virginia Medical School Biomarker Discovery Laboratory

## Investigators

John Semmes, Ph.D.  
John Davis, M.D.  
Jose Diaz, M.D., Ph.D.  
Rick Drake, Ph.D.  
Christine Laronga, M.D.  
Paul Schellhammer, M.D.  
Jeffery T. Wadsworth, M.D.

## Fellows

Alberto Corica, M.D.  
Daniel Holterman, Ph.D.  
Gunjan Malik, Ph.D.  
Lining Qi, Ph.D.

## Staff

Diane Brassil  
Lisa Cazares  
MaryAnn Clements  
Tarek Kendil  
Brian Main  
Michelle Moody  
Michael Ward

E  
V  
M  
S



## Biostatistics/Computation

WMRI

William Cooke, Ph.D.  
Dasha Malyrenko, Ph.D.  
Denis Manos, Ph.D.  
Michael Trossett, Ph.D.  
Eugene Tracy, Ph.D.